Detection of Failure Rate Increases

G. Lorden
California Institute of Technology

I. Eisenberger
Communications Systems Research Section

The problem of devising systematic policies for replacement of equipment subject to wear-out involves the detection of increases in failure rates. Detection procedures are defined as stopping times N with respect to the observed sequence of random failures. The concepts of "quickness of detection" and "frequency of false reactions" are made precise and a class of procedures is studied which optimizes the former asymptotically as the latter is reduced to zero. Results of Monte Carlo experiments are given which show that efficient quickness of detection is attainable simultaneously for various levels of increase in failure rates.

I. Introduction

The present formulation of the problem of detecting failure rate increases arose in the study of replacement policies for equipment which may possibly be subject to wear-out, under the assumption that little is known *a priori* about when the onset of wear-out is likely to occur, or even whether it will occur. The desired type of policy is a rule utilizing failure data themselves to determine that the failure rate has increased. When such determination has occurred, some previously specified action is taken, e.g., investigation of causes or ordering of replacements. It is desired that this action be taken as soon as possible after a specified level of increase in the failure rate has occurred, and it is by no means necessary to estimate when that increase began. Thus, in mathematical terms, the kind of statistical procedure sought is a stopping time N for an observed sequence of random variables $X_1, X_2, X_3, \cdots$. That is, N is a random variable with possible values $1, 2, \cdots$, and ∞ (i.e., never stops), such that for every $n = 1, 2, \cdots$, the event $\{N = n\}$ depends on $X_1, \cdots, X_n$ only. The X_i's are times between successive failures, and are assumed to be independent, with exponential densities

$$f_{\lambda_i}(x) = \begin{cases} \lambda_i e^{-\lambda_i x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \qquad \lambda_i > 0, i = i, 2, \cdots \tag{1}$$

In order to define a simple criterion for quickness of reaction to increases in the failure rate, it is convenient to consider the following situation: For some $m = 1, 2, \cdots$,

$$\lambda_1 = \lambda_2 = \cdots = \lambda_{m-1} = \lambda \text{ (known)}$$

and

$$\lambda_m = \lambda_{m+1} = \cdots = (1+\theta)\lambda, \theta > 0 \tag{2}$$

Note that Eq. (2) specifies that the increase in failure rate from λ to $(1+\theta)\lambda$ occurs instantaneously after X_{m-1} is observed. Denote by P_m and E_m probabilities and expectations for $m = 1, 2, \cdots$, and denote the same by P_0 and E_0 when $\lambda = \lambda_1 = \lambda_2 = \cdots$. A reasonable measure of quickness of detection of increases occurring at time m is the smallest number C_m such that

$$E_m[N-(m-1)|X_1 = x_1, \cdots, X_{m-1} = x_{m-1}] \leqq C_m$$

for all $x_1, \cdots, x_{m-1}$ such that $N \geqq m$. As a kind of "worst case" criterion, define $\bar{E}_\theta N$ as the largest of the C_m's, i.e.,

$$\bar{E}_\theta N = \sup_{m \geqq 1} C_m \tag{3}$$

The desire to have small $\bar{E}_\theta N$ for $\theta > 0$ must, of course, be balanced against the need to have a controlled frequency of "false reactions." In other words, when there is no increase in failure rate, then N should be large, hopefully infinite. It is shown in Ref. 1, however, that in order to have $\bar{E}_\theta N$ finite for some $\theta > 0$ it is necessary that N have finite expectation even under P_0. An appropriate type of restriction on false reactions, therefore, is

$$E_0 N \geqq \gamma > 1 \tag{4}$$

where γ is to be prescribed.

The problem under investigation can now be formulated more precisely. Among all stopping times N satisfying Eq. (4) for prescribed γ, determine one which minimizes (or nearly minimizes) $\bar{E}_\theta N$ over a specified range, $\theta_1 \leqq \theta \leqq \theta_2$. In Ref. 1, it is shown that as $\gamma \to \infty$ the minimum possible $\bar{E}_\theta N$ ($\theta > 0$ fixed) is asymptotic to

$$\frac{\log \gamma}{\log(1+\theta) - \dfrac{\theta}{1+\theta}} \tag{5}$$

where the denominator is the Kullback–Leibler information number when $(1+\theta)\lambda$ is true and the alternative is λ. In that paper, it is also demonstrated that a "maximum likelihood" procedure, $\hat{N}$, achieves the asymptotic minimum simultaneously for all $\theta > 0$. (The rate of approach to the asymptotic minimum depends on θ, however.) These procedures are defined for the present case of exponential distributions in *Section III* and computationally simpler modifications are introduced, along with Monte Carlo results. It is helpful to take up first the case of a single alternative $\theta > 0$, which will be done in *Section II. Section IV* treats the case where λ is unknown.

II. Simple Alternative

Motivated by the problem of control charts in quality control, E. S. Page (Ref. 2) proposed a general procedure for detecting a change from one density to another at an unknown location in a sequence of random variables. His procedure consists of repeated applications of a sequential probability ratio test (SPRT) which in the present context is definable by the inequalities (for fixed $\theta > 0$)

$$0 < n \log(1+\theta) - \theta S_n < \log \gamma \tag{6}$$

where $S_n = X_1 + \cdots + X_n$, γ is chosen > 1, and it is assumed from this point on that $\lambda = 1$ (which can always be achieved by scaling the X's). The procedure is to stop as soon as the right-hand inequality is violated, with the proviso that if the left-hand inequality is violated first all observations up to that point will be discarded and the procedure "recycled," with $S_1, S_2, \cdots$, denoting cumulative sums of the new observations.

The following equivalent formulation is convenient to apply: stop the first time that

$$T_n \geqq \log \gamma \tag{7}$$

where $T_0 = 0$ and for $n = 1, 2, \cdots$,

$$T_n = \max(0, T_{n-1} + \log(1+\theta) - \theta X_n) \tag{8}$$

It is illuminating also to view Page's procedure in another way. Stopping occurs when *for some* $k \geqq 1$ the last k observations, $X_{n-k+1}, \cdots, X_n$, are "significant" in the sense of a one-sided SPRT, i.e.,

$$k \log(1+\theta) - \theta(X_{n-k+1} + \cdots + X_n) \geqq \log \gamma$$

Let α, $1-\beta$, respectively, denote the probabilities under P_0, P_1 that the procedure stops before recycling. Then the expected number of cycles is evidently α^{-1}, $(1-\beta)^{-1}$, respectively. If N_1 denotes the number of observations required to violate either inequality, then by Wald's equation (Ref. 3) for the expected value of the sum of a random number of independent and identically distributed variables, the number N of observations taken by Page's procedure satisfies

$$E_0 N = \alpha^{-1} E_0 N_1 \tag{9}$$

and

$$E_\theta N = (1-\beta)^{-1} E_\theta N_1 \tag{10}$$

Furthermore,

$$\overline{E}_\theta N = (1-\beta)^{-1} E_\theta N_1 \tag{11}$$

since obviously $\overline{E}_\theta N \geqq E_\theta N$. And $\overline{E}_\theta N \leqq E_\theta N$ by the following argument. Observing $X_1 = x_1, \cdots, X_{m-1} = x_{m-1}$ determines that $T_{m-1} = t \geqq 0$ (depending on $x_1, \cdots, x_{m-1}$). Since $X_m, X_{m+1}, \cdots$, are independent of past X's, the sequence $T_m, T_{m+1}, \cdots$, behaves just as $T_1, T_2, \cdots$, would if one started with $T_0 = t \geqq 0$. Since this last obviously would not make any succeeding T's smaller, it would not increase the time required to reach $\log\gamma$. This proves Eq. (11).

Since $\alpha \leqq \gamma^{-1}$ by the usual estimates of SPRT error probabilities (Ref. 4), evidently

$$E_0 N \geqq \gamma\, E_0 N_1 \geqq \gamma$$

Furthermore $(1-\beta)^{-1} E_\theta N_1$ is asymptotic to $\log\gamma$ divided by the information number, by virtue of the usual Wald formulas for expected sample sizes. Thus Page's procedure does approach asymptotically the minimum $\overline{E}_\theta N$ (Expression 5).

Using Eqs. (9) and (11) one can obtain good approximations to $\overline{E}_0 N$ and $E_\theta N$ in terms of γ from the approximations of SPRT error probabilities and expected sample sizes for exponential densities given in Ref. 3. For the boundaries 0 and $\log\gamma$ in Expression (6), these approximations are as follows:

$$1-\beta = \alpha\gamma G(\theta) = \frac{\theta\gamma G(\theta)}{\gamma G(\theta)(1+\theta) - 1} \tag{12}$$

where

$$G(\theta) = \frac{(1+\theta)\log(1+\theta) - \theta}{\theta - \log(1+\theta)}$$

$$(\log(1+\theta) - \theta)\, E_0 N_1 = \alpha \log(\gamma(1+\theta)) - (1-\alpha)\theta \tag{13}$$

$$\left(\log(1+\theta) - \frac{\theta}{1+\theta}\right) E_\theta N_1 = (1-\beta) \times \left(\log\gamma + \frac{\frac{1}{2}(1+\theta)(\log(1+\theta))^2}{(1+\theta)\log(1+\theta) - \theta}\right) - \frac{\beta\theta}{1+\theta} \tag{14}$$

The approximations (12)–(14) give approximations to $\overline{E}_0 N$ and $E_\theta N$ by Eqs. (9) and (11). The accuracy of these approximations is indicated by the following comparison (Table 1) with the values based on the exact formulas in Ref. 5 (which entail considerably more calculation).

III. Composite Alternative

For the problem of minimizing $\overline{E}_\theta N$ over a range $\theta_1 \leqq \theta \leqq \theta_2$ subject to $E_0 N \geqq \gamma$, it is natural to consider simultaneous Page procedures. Performing Page's procedures simultaneously for all alternatives $\theta \epsilon [\theta_1, \theta_2]$ results in stopping when for some $k \geqq 1$ the last k observations satisfy

$$\max_{\theta_1 \leqq \theta \leqq \theta_2} [k \log(1+\theta) - \theta(X_{n-k+1} + \cdots + X_n)] \geqq \log\gamma$$

This rule is computable since the indicated maximum is attained either at θ_1, or θ_2, or at the maximum likelihood estimate, given by $\hat{\theta} = (X_{n-k+1} + \cdots + X_n)k^{-1} - 1$. This is the procedure, $\hat{N}$, which achieves the asymptotic minimum (Expression 5) for every $\theta \epsilon (\theta_1, \theta_2)$, as shown in Ref. 1. In that paper the computation of this type of procedure is discussed.

The results of preliminary Monte Carlo experiments indicated that in the "small sample case," i.e., $E_0 N \leqq 2000$, when θ_2/θ_1 is not very large, the improvement of $\overline{E}_\theta N$ for $\theta \epsilon [\theta_1, \theta_2]$ achieved by $\hat{N}$ in comparison to Page's procedure is already achieved to a large extent by the simpler rule which uses two simultaneous Page procedures, one for each of the alternatives θ_1, θ_2. Accordingly, the following results are limited to this dual-Page procedure, $\tilde{N}$. Extensive Monte Carlo sampling was carried out with $\theta_1 = 0.5$ and $\theta_2 = 0.8$. Thus, the range of alternatives where efficient performance was most emphasized represented 50% to 80% increases in failure rate. The values $\gamma = 60$ and $\gamma = 100$ were chosen, resulting in estimates of $E_0\tilde{N}$ equal to 508 and 936, respectively. The results are summarized in Table 2. (The tolerances given are sample variances.) Just as for a single Page procedure, $\overline{E}_\theta \tilde{N} = E_\theta \tilde{N}$ for $\theta > 0$.

The value $\theta = -0.1$ is included in Table 2 to indicate how $\tilde{N}$ performs if the true failure rate remains 10% less than the nominal value. In both cases $\gamma = 60$, 100, the frequency of false reactions is about one-third as large as when the failure rate equals the nominal value.

Note that the efficiency of $\tilde{N}$ is about 96% and 98%, respectively, for $\gamma = 60$, 100, and θ between 0.5 and 0.8 (the efficiency estimate of 100.1% resulting from sampling error). For θ outside the chosen interval [0.5, 0.8], the efficiency falls off gradually but is still quite high between 0.4 and 1.0, particularly for the smaller γ.

Comparison of the results for $\gamma = 60$ and 100 indicates that a much larger $E_0\tilde{N}$ is obtainable for a relatively small increase in $\overline{E}_\theta\tilde{N}$'s. An increase of about 15% in $E_\theta\tilde{N}$'s between the two cases yields nearly a doubling of $E_0\tilde{N}$.

For fixed γ, there is a convenient rule of thumb that fairly well approximates $E_\theta\tilde{N}$ over the indicated range; namely, $E_\theta\tilde{N}$ is inversely proportional to θ (or, equivalently, the percent increase in the failure rate). Table 3 indicates the accuracy of the approximation $\theta E_\theta\tilde{N} =$ constant in the case of the Monte Carlo results of Table 2. The rule of thumb exhibits a similar degree of accuracy in approximating the $E_\theta N$ (from Eqs. 10 and 14) of a Page procedure for θ with γ chosen (depending on θ) to achieve a prescribed E_0N (from Eqs. 9 and 13).

Having chosen θ_1, θ_2 for a dual-Page procedure, the problem naturally arises of how to select γ to achieve a prescribed E_0N. (The corresponding problem for a single-Page procedure is solvable by successive approximations using Eqs. 9 and 13.) Unfortunately, it seems to be very difficult to derive approximations for $E_0\tilde{N}$ in terms of γ. Bounds are obtainable, however, from the following simple considerations. In the case $\gamma = 60$, for example, the Page procedures for $\theta_1 = 0.5$ and $\theta_2 = 0.8$ have frequencies of false reaction, $1/E_0N$, equal to 1/588 and 1/1026, respectively, according to Eqs. (9) and (13). Evidently, the dual procedure $\tilde{N}$ has frequency of false reaction at least 1/588 and at most $1/1026 + 1/588 = 1/374$. Thus, $374 < E_0\tilde{N} < 588$. Note that the Monte Carlo result of 508 is in fact closer to the upper bound, which is also true in the case $\gamma = 100$.

It is not very difficult to estimate $E_0\tilde{N}$ by Monte Carlo methods accurately enough to choose γ, once the range has been narrowed by using the bounds just described. Since the values of $\overline{E}_\theta\tilde{N}$ increase rather slowly compared to $E_0\tilde{N}$ as γ is made larger, there is little harm in choosing γ conservatively.

The single- and dual-Page procedures, N and $\tilde{N}$, and the maximum likelihood procedure $\hat{N}$ all have a pleasant property: when $\theta = 0$, the time to stop is approximately exponentially distributed. To see this for N, note that the cycles defined by Expression (6) are a sequence of Bernoulli trials, and stopping occurs upon the first failure to recycle (i.e., violation of the right-hand inequality). Thus, the number of cycles is geometrically distributed, and when γ (and hence the number of cycles) is large, the number of observations also is nearly geometrical (approximately exponential) in distribution. The same holds true for $\tilde{N}$ and $\hat{N}$, since recycling of the Page procedure for θ_1 entails (Ref. 1) recycling of the Page procedures for all $\theta > \theta_1$, i.e., the initial conditions are duplicated. The approximate exponential distribution gives a reasonable indication of the probabilities of "unlucky" early false reactions.

IV. The Case of Unknown λ

By dealing with the sequence of ratios $S_2/S_1, S_3/S_2, \cdots$, one can obviously develop procedures whose performance does not depend on the scale factor, λ. The sequence $\{S_n\}$ is a Poisson process (so long as the failure rate remains constant) and it is well known (Ref. 6) that the conditional distribution of S_n given $S_{n+1} = t$ is the same as the distribution of the largest of n independent variables uniformly distributed on $[0, t]$. Thus,

$$P(S_n \leq x \mid S_{n+1} = t) = \begin{cases} \left(\dfrac{x}{t}\right)^n, & 0 \leq x < t \\ 1 \quad , & x \geq t \end{cases}$$

and hence

$$P\left(\left(\frac{S_n}{S_{n+1}}\right)^n \leq u \mid S_{n+1} = t\right) = P\left(S_n \leq u^{\frac{1}{n}} t \mid S_{n+1} = t\right)$$
$$= \begin{cases} u, & 0 \leq u \leq 1 \\ 1, & u \geq 1 \end{cases} \qquad (15)$$

Since this last expression doesn't depend on t, evidently $(S_n/S_{n+1})^n$ is uniformly distributed on $(0, 1)$ and independent of S_{n+1}. In fact, $(S_n/S_{n+1})^n$ is independent of $S_{n+1}, \cdots, S_m$ (jointly) for any $m > n + 1$, since the conditional distributions above are unchanged if the condition $S_{n+1} = t$ is augmented by specifying $S_{n+2} = t_{n+2}, \cdots, S_m = t_m$. Therefore, $(S_n/S_{n+1})^n$ is evidently independent of $(S_{n+1}/S_{n+2})^{n+1}, \cdots, (S_{m-1}/S_m)^{m-1}$ jointly. For fixed m, the last statement is true for all $n < m - 1$, and hence

$$S_1/S_2, (S_2/S_3)^2, \cdots, (S_{m-1}/S_m)^{m-1}$$

are mutually independent (for every $m \geq 3$). Thus, the random variables in the infinite sequence

$$S_1/S_2, (S_2/S_3)^2, (S_3/S_4)^3, \cdots$$

are independent and (by the remark following Eq. 15) each is uniformly distributed on (0, 1). It is easy to verify that if U is uniformly distributed on (0, 1), then $\log U^{-1}$ is exponentially distributed with mean 1. Thus,

$$\log \frac{S_2}{S_1}, 2 \log \frac{S_3}{S_2}, 3 \log \frac{S_4}{S_3}, \cdots \tag{16}$$

are independent and exponentially distributed with mean one, *regardless of the true value of* λ. The (single or dual) Page procedures of the preceding sections, when applied to the sequence (16), will therefore yield the same E_0N as before.

Under what circumstances will the sequence (16) be independent and exponentially distributed with mean $1/(1+\theta)$? Obviously, it suffices that $S_1, S_2, \cdots$ have the same distribution as $W_1^{1/1+\theta}, W_2^{1/1+\theta}, \cdots$, where $\{W_n\}$ is a Poisson process. This is the case if, for example, the S_n's are the times of successive failures occurring in a family of repairable parts under the following assumptions. Their failure rate functions depend only on age (the effects of previous failures disappearing upon repair) and are Weibull with shape parameter $\alpha = 1/(1+\theta)$ and arbitrary scale parameters (not necessarily the same for all parts).

The behavior of sequence (16) when the failure rate changes abruptly at time m can be described approximately by noting that

$$n \log \frac{S_{n+1}}{S_n} = \log \left(1 + \frac{X_{n+1}}{nX_n}\right)^n \approx \frac{X_{n+1}}{\overline{X}_n} \text{ for large } n \tag{17}$$

where $\overline{X}_n = S_n/n$. If the failure rate is λ for $X_1, \cdots, X_m$, then changing it to $(1+\theta)\lambda$ thereafter multiplies $X_{m+1}, X_{m+2}, \cdots$ by $1/(1+\theta)$, while $\overline{X}_n$ is largely unaffected so long as $n - m << m$. For $n >> m$, however, the contribution of $X_1, \cdots, X_m$ to $\overline{X}_n$ becomes small and $n \log (S_{n+1}/S_n)$ begins to approach an exponential distribution with mean one again. If the failure rate changes after X_m from a constant to a Weilbull failure rate function with $\alpha = 1/(1+\theta)$ (keeping the same scale parameter), then it is easy to see that for $n >> m$ the variables $n \log (S_{n+1}/S_n)$ will be approximately independent exponential with mean $1/(1+\theta)$.

In summary, then, the application of the procedures studied in the preceding sections to the sequence (16) leaves E_0N unchanged and should result in efficient detection whenever the failure rate increases sharply and continues to increase in the form of a Weibull failure rate function.

References

1. Lorden, G., "Procedures for Reacting to a Change in Distribution" (submitted to *Ann. Math. Statist.*).
2. Page, E. S., "Continuous Inspection Schemes," *Biometrika*, Vol. 41, pp. 100–115, 1954.
3. Lorden, G., "Sequential Tests for Exponential Distributions" (submitted to *Ann. Math. Statist.*).
4. Wald, A., *Sequential Analysis*. John Wiley & Sons, Inc., New York, 1947.
5. Dvoretzky, A., Kiefer, J., and Wolfowitz, J., "Sequential Decision Problems for Processes with Continuous Time Parameter: Testing Hypotheses," *Ann. Math. Statist.*, Vol. 24, pp. 254–264, 1950.
6. Doob, J. L., *Stochastic Processes*. John Wiley & Sons, Inc., New York, 1953.

Table 1. Comparison of actual and approximate expected stopping times

θ	γ	E_0N		$E_\theta N$	
		Actual	Approximate	Actual	Approximate
0.4	20	422.1	418.5	47.9	47.8
0.6	50	676.0	673.6	36.4	36.4
0.9	40	342.0	340.3	20.2	20.2

Table 2. Number of observations before detection (Monte Carlo sampling)

	Value of θ								
	−0.1	0	0.25	0.4	0.5	0.6	0.8	1.0	1.5
$E_\theta\tilde{N}$ (60)	1701	508	96.3	53.9	42.2	35.0	26.2	21.3	15.8
	±181	±21	±2.5	±1.2	±1.0	±0.7	±0.4	±0.3	±0.2
% Efficiency[a]			85.5	94.7	96.1	96.2	96.3	95.6	89.1
$E_\theta\tilde{N}$ (100)	2756	936	128.0	69.0	48.3	40.5	30.2	24.5	17.6
	±236	±48	±3.5	±1.6	±1.0	±0.7	±0.5	±0.3	±0.2
% Efficiency[a]			81.5	89.8	100.1	98.1	97.2	95.6	91.1

[a]The efficiency was estimated using the ratio of (sampled) $E_\theta\tilde{N}$ to the $E_\theta N$ of a Page procedure for θ having the same E_0N as $E_0\tilde{N}$ (sampled).

Table 3. Values of $\theta E_\theta\tilde{N}$ (sampled)

γ	Value of θ						
	0.25	0.4	0.5	0.6	0.8	1.0	1.5
60	24.1	21.6	21.1	21.0	21.0	21.3	23.7
100	32.0	27.6	24.2	24.3	24.2	24.5	26.4